

# **Projet Postdoctoral BLRI**

**Jean-Baptiste Bernard**

## **Nom du Projet :**

Implémentation, validation expérimentale et valorisation d'un modèle Bayésien de prédiction des comportements oculomoteurs pendant la lecture.

## **Nom du Postdoctorant :**

**Jean-Baptiste Bernard**, Postdoctorant Laboratoire de Psychologie Cognitive (LPC)

## **Encadrement BLRI :**

**Eric Castet**, Laboratoire de Psychologie Cognitive (LPC).

**Benoit Favre & Nuria Gala**, Laboratoire Informatique Fondamentale de Marseille (LIF).

**George Linares**, Laboratoire d'Informatique d'Avignon (LIA).

## Résumé :

Ce projet vise à implémenter et à valider par des expériences comportementales un modèle permettant de rendre compte des processus visuels et cognitifs se déroulant lors de la lecture d'un texte. Contrairement aux modèles de lecture actuels les plus aboutis, ce modèle prend en compte les limitations visuelles de bas niveau propres à chaque lecteur, ainsi que la composition orthographique des mots identifiés pendant la lecture d'un texte. Le modèle est basé sur la théorie de l'information et adopte une représentation probabiliste de la présence des différentes lettres et mots du texte. Ces probabilités sont mises à jour à chaque nouvelle fixation ( $n$ ) en les intégrant au fur et à mesure avec les informations des fixations précédentes ( $n-i, i \geq 1$ ). La position de la prochaine fixation ( $n + 1$ ) est choisie de façon à maximiser le gain d'information apporté par cette nouvelle fixation. L'implémentation informatique de ce modèle permettra de produire une séquence oculomotrice et une vitesse de lecture spécifique en réponse à n'importe quel texte. Ce modèle sera testé avec des sujets normo-lecteurs pour prédire l'influence (1) des limites visuelles, (2) des limites oculomotrices et (3) des processus linguistiques (propriétés lexicales, sémantiques et syntaxiques du texte) sur le comportement oculomoteur et les performances du lecteur. Cette dissociation expérimentale des facteurs limitants est rendue possible par un effet de la fréquence propre des mots ou de leur prédictibilité sur les probabilités de reconnaissance des mots. Une fois le modèle validé, l'analyse du comportement oculomoteur de sujets ayant des performances de lecture réduites (sujets malvoyants ou dyslexiques) permettra de déterminer quels facteurs du modèle (visuel, oculomoteur ou linguistique) peuvent expliquer les déficits des lecteurs, et quelles tâches de rééducation seraient les plus adaptées aux sujets. Les études menées pendant ce projet permettront également le développement d'un test hautement standardisé de lecture.

# Description du projet :

## Introduction

La lecture de texte est une activité complexe faisant appel à différents processus qui interagissent depuis les étapes visuelles de bas niveau (la reconnaissance des traits visuels des lettres) jusqu'aux processus cognitifs de très haut niveau (la compréhension du texte présenté). L'extraction d'information visuelle est basée sur la position du regard (et plus précisément de la fovéa) lors de chaque fixation oculaire sur les mots d'un texte. Les performances visuelles de reconnaissance de lettres et de mots diminuent en effet rapidement lorsque la distance à la fovéa augmente, essentiellement à cause des effets dus à l'acuité visuelle et au phénomène d'encombrement [1]. Ainsi, peu de lettres peuvent être reconnues avec certitude lors d'une fixation. Cette limitation visuelle est compensée par le fait qu'un traitement para-fovéal est généralement déjà effectué (à droite de la fixation pour la lecture de gauche à droite), ainsi que l'utilisation d'une inférence lexicale ou contextuelle. L'information interagit donc à différents niveaux (niveaux lettre, mot et phrase à travers des processus visuels et cognitifs mis en place une ou plusieurs fixations en amont) pour permettre l'identification et l'intégration sémantique des mots d'un texte. Les modèles les plus évolués qui prédisent la performance et les mouvements des yeux lors de la lecture sont les modèles EZ Reader et SWIFT [2]. Ces modèles prennent en compte les limitations visuelles de façon sommaire. Ils définissent la distance du centre du mot à la fovéa comme un indice modulant le temps de fixation sur un mot sans étudier les variations de performance de reconnaissance de lettres dans le champ visuel des lecteurs. Chaque mot est ainsi représenté uniquement par sa taille et sa fréquence lexicale, ce qui élimine l'aspect orthographique lors de la reconnaissance des mots dans ces modèles, un effet pourtant fondamental lors du traitement lexical.

## Définition du modèle

Dans ce projet, nous souhaitons développer et implémenter un modèle de lecture plus complet que les modèles existant, traitant à la fois les aspects visuels, oculomoteurs et cognitifs de la lecture. Il s'agit d'un modèle probabiliste, inspiré du modèle sommaire de lecture « Mr Chips » pour malvoyants [3], qui est basé sur les probabilités d'identification de lettres et de mots présents dans une phrase<sup>1</sup>. Chaque mot du texte est traité en évaluant quelle est sa probabilité d'être un des mots du lexique mental du lecteur. Une particularité de ce modèle est l'utilisation d'un « profil d'empan visuel » (PEV, [5]) pour réduire l'incertitude associée aux lettres et aux mots à chaque fixation (Figure 1). Ces PEV sont mesurées psychophysiquement pour différentes durées et pour chaque lecteur. Ils permettent de déterminer l'information visuelle extraite par le lecteur pour chaque lettre en fonction de la position de la lettre dans le champ visuel du lecteur. Une inférence Bayésienne permet alors de combiner l'information acquise lors de la fixation  $n$  avec l'information acquise lors des fixations précédentes, mettant ainsi à jour les probabilités d'activation des mots du lexique (Figure 2). La position de la prochaine fixation est calculée par le modèle en maximisant le gain d'information apporté par cette fixation. Cette hypothèse est en accord avec certains modèles Bayésiens de la vision qui indiquent que les mouvements des yeux dans des tâches de recherche visuelle peuvent être prédits par le gain d'information apporté par chaque fixation à venir [6]. Un autre but majeur de ce projet est d'intégrer certains facteurs linguistiques en tenant compte quantitativement de l'influence de ces facteurs sur les probabilités d'identification des mots. Par exemple, la probabilité d'identification d'un mot du lexique est directement liée à sa fréquence d'occurrence (un mot fréquent aura une probabilité d'identification plus élevée) ou à la composition de la phrase (un mot prédictible par la syntaxe ou la sémantique de la phrase aura une probabilité d'identification plus élevée). D'autre part, la durée de chaque fixation sera calculée en fonction de la quantité d'information extraite à chaque fixation. Une fixation « peu informative » (sur un mot fréquent et très prédictible) nécessitera ainsi un temps de fixation moins élevé qu'une fixation « très informative » (sur un mot peu fréquent et peu prédictible). La notion de prédictibilité est donc une notion clé dans ce modèle pour prendre en compte les

---

<sup>1</sup> Il serait également possible d'incorporer dans un second temps des unités intermédiaires de groupes de lettres comme des graphèmes ou des phonèmes, mais la première version du modèle sera basée sur une étape de reconnaissance des lettres, précédant l'étape de reconnaissance de mots, la reconnaissance de lettres apparaissant comme l'unité de base dans la reconnaissance de mot et la lecture [4].

processus cognitifs menant à l'identification d'un mot. Une partie importante de notre projet sera ainsi destinée à définir une ou plusieurs métriques de prédictibilité d'un mot basées sur des mots, des groupes de mots voire des phrases précédemment lus.

### **Implémentation et validation du modèle**

J'ai récemment implémenté en Matlab une première version du modèle pour rendre compte dans un premier temps de la position des fixations, sans prendre en compte les mesures de prédictibilité et de durées des fixations [7] (voir Figure 3 pour un exemple de lecture d'une phrase par le modèle). Cette version utilise un lexique composé des 30,000 mots français les plus fréquents. Elle apparaît très prometteuse car elle permet de mettre en évidence certaines propriétés oculomotrices connues de la lecture (position préférée du regard sur les mots, proportion de mots sautée, re-fixation de mots longs). La première étape du projet consistera en l'implémentation d'une nouvelle version du modèle qui incorporera les durées de fixation. Parallèlement, j'explorerai les différentes métriques possibles pour caractériser la prédictibilité des mots. Plusieurs métriques devront être définies qui prédisent la probabilité de présence d'un mot  $n$  en se basant sur les mots précédents  $(n-i)_{1 \leq i \leq N}$ , avec possibilité de prendre en compte le nombre de mots  $N$  précédents. Il sera également possible de déterminer l'influence de ces mots en fonction de leur distance spatiale ou temporelle au mot  $n$ . Ces méthodes de prédiction permettront de rendre compte des propriétés syntaxiques (par exemple sous-catégorisation verbale) ou sémantiques (mots co-occurents, sémantiquement ou thématiquement proches) des phrases. Plusieurs méthodes basées sur des mesures statistiques du langage écrit seront ainsi testées. La prédiction pourrait ainsi se faire à court terme (typiquement par des modèles  $n$ -grams ou Word-Embedding [8]), à plus long terme pour capturer les relations sémantiques entre les mots (Latent Semantic Analysis ou Latent Dirichlet Analysis [9]), voire à partir de modèles de séquence (réseaux récurrents LSTM (Long Short Term Memory) [10] pour prendre en compte le maximum d'informations concernant la totalité des mots et des phrases précédant la fixation. Afin de valider ce modèle, je réaliserai au moins deux expériences principales de psychophysique. La première expérience consistera en la lecture de mots n'ayant aucun lien entre eux (pas de liens lexicaux ou sémantiques). La fréquence des mots constituant la phrase sera contrôlée, et les PEV seront mesurés pour chaque sujet participant à l'expérience (au moins 20 sujets). Les mouvements des yeux des sujets seront mesurés à l'aide de l'oculomètre Eyelink 1000 que j'utilise depuis plusieurs années. Les coefficients optimaux du modèle seront alors déterminés pour chaque sujet (coefficients avec le maximum de vraisemblance) de façon à tenir compte de l'empan visuel, de la longueur, de la fréquence d'occurrence et de la composition orthographique des mots. La deuxième expérience sera similaire avec cette fois de vraies phrases dont les dépendances entre les mots seront quantifiées. Plusieurs mesures de prédictibilité étudiées dans la première étape seront testées de façon à définir la métrique optimale pour interpréter les données expérimentales.

### **Application pour les sujets avec performance de lecture réduite**

Dans un deuxième temps, nous souhaitons tester le modèle sur des sujets avec perte de vision centrale simulée (simulation de la Dégénérescence Maculaire Liée à l'âge, [11,12]). En effet, une grande force de ce modèle est qu'il peut également être appliqué à des sujets sans vision centrale en prédisant des séquences oculomotrices en 2 dimensions (la fovéa ne peut être posée sur le texte à lire dans ce cas précis). Aucun autre modèle de lecture n'en est capable à ce jour. La modification principale interviendra au niveau des caractéristiques du PEV, certaines positions du champ visuel ne permettant pas la reconnaissance de lettres. Déterminer la part des processus visuels, oculomoteurs et linguistiques dans les difficultés de lecture des sujets avec perte de vision centrale permettra d'apporter des éléments de réponse importants à une question non résolue dans les difficultés de lecture des patients sans vision centrale où la part oculomotrice est souvent négligée [13]. Les performances du modèle permettront aussi de diriger les efforts de réhabilitation chez les patients basse vision, par exemple vers un apprentissage perceptif dans le cas de limitations visuelles, ou vers un apprentissage oculomoteur dans le cas de limitations oculomotrices. Une autre perspective intéressante à long terme est de généraliser les prédictions de mon modèle à la lecture des sujets dyslexiques. Ces expériences ne pourraient pas être raisonnablement effectuées pendant les 2 ans du projet, mais il est intéressant de noter que les principes du modèle s'appliquent aux spécificités de la dyslexie [14].

## Développement d'une base de phrases calibrées pour la mesure des vitesses de lecture

Dans un troisième temps, je souhaiterais utiliser les mesures de prédictibilité validées dans les étapes précédentes pour concevoir une base de phrases strictement contrôlée pour la mesure des vitesses de lecture chez les lecteurs avec vitesse de lecture normale ou réduite (sujets dyslexiques ou malvoyants). Les bases de phrases disponibles actuellement (y compris celles que j'ai développée et utilisée [11]) prennent rarement en compte ce facteur de prédictibilité qui est pourtant à l'origine d'une grande variabilité des mesures de vitesses de lecture [15]. Un outil informatique sera développé pour automatiser la sélection de phrases strictement calibrées à partir de livres numériques [11]. La définition et la validation de cette base de phrases permettrait sa mise à disposition et son utilisation à l'ensemble de la communauté scientifique. Une version en langue Anglaise pourrait également être développée<sup>2</sup>. Dans le futur, la mesure des mouvements des yeux de lecteurs lisant les phrases de cette base pourrait permettre de caractériser les limites perceptives, oculomotrices et linguistiques des sujets. Utiliser notre modèle sur ces mouvements des yeux pourrait alors permettre de déterminer les différents facteurs limitant les performances de lecture des sujets.

### Calendrier

**E1:** Implémentation du modèle

**E2:** Exploration des différentes métriques de prédictibilité

**E3:** Expériences psychophysiques/oculomotrices de validation (lecture normale)

**E4:** Expériences psychophysiques/oculomotrices de validation (lecture avec perte de vision centrale simulée)

**E5:** Développement d'un test de lecture

**A1:** Article scientifique 1 (ou 1a et 1b): Définition et Validation du modèle de lecture

**A2:** Article scientifique 2: Utilisation du modèle pour caractériser les baisses de vitesse de lecture lors de la lecture avec perte de vision centrale

**A3:** Article scientifique 3: Test de lecture

		Semestres			
		S1	S2	S3	S4
Tâches	E1				
	E2				
	E3				
	E4				
	E5				
	A1				
	A2				
	A3				

### Références

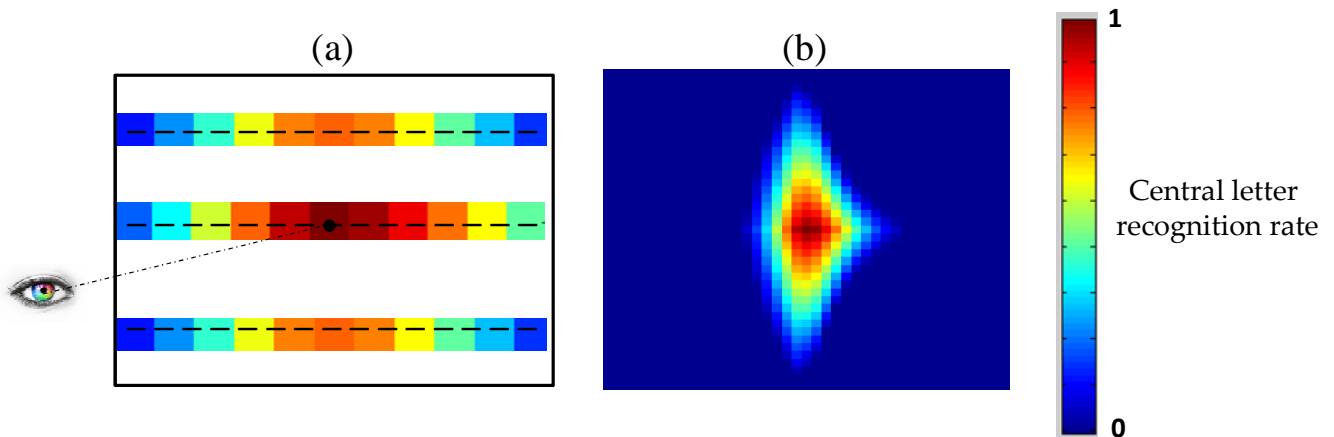
- Pelli DG, Palomares M, Majaj NJ. Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis.* 2004;4: 1136–1169.
- Reichle ED, Rayner K, Pollatsek A. The E-Z reader model of eye-movement control in reading: comparisons to other models. *Behav Brain Sci.* 2003;26: 445–476; discussion 477–526.
- Legge GE, Legge GE, Klitz TS, Klitz TS, Tjan BS, Tjan BS, et al. Mr. Chips: An ideal-observer model of reading. *Psychol Rev.* 1997;104: 524–553.
- Pelli DG, Farell B, Moore DC. The remarkable inefficiency of word recognition. *Nature.* 2003;423: 752–756.
- Legge GE, Mansfield JS, Chung ST. Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Res.* 2001;41: 725–743.
- Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. *Nature.* 2005;434: 387–391. doi:10.1038/nature03390
- Jean-Baptiste Bernard. An information-based model to predict oculomotor behaviors in reading. International Workshop about Visual Aspects of Reading. Grasmere, England; 2015.
- Mikolov T, Yih W, Zweig G. Linguistic regularities in continuous spaceword representations. In HLT-NAACL. 2013.
- Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. *J Mach Learn Res.* 2003;3: 993–1022.
- Hochreiter S, Schmidhuber J. Long Short-term Memory. 1995.
- Bernard J-B, Anne-Catherine S, Eric C. Page mode reading with simulated scotomas: A modest effect of interline spacing on reading speed. *Vision Res.* 2007;47: 3447–3459.
- Aguilar C, Castet E. Gaze-contingent simulation of retinopathy: some potential pitfalls and remedies. *Vision Res.* 2011;51: 997–1012.
- Calabrèse A, Bernard J-B, Faure G, Hoffart L, Castet E. Eye Movements and Reading Speed in Macular Disease: The Shrinking Perceptual Span Hypothesis Requires and Is Supported by a Mediation Analysis. *Invest Ophthalmol Vis Sci.* 2014;55: 3638–3645.
- Gori S, Facoetti A. How the visual aspects can be crucial in reading acquisition? The intriguing case of crowding and developmental dyslexia. *J Vis.* 2015;15: 15.1.8.
- Rubin GS. Measuring reading performance. *Vision Res.* 2013;90: 43–51.

<sup>2</sup> D'un point de vue linguistique, il faudra bien entendu prendre en compte les différences au niveau du lexique et des structures de phrases différentes en langue anglaise.

## Annexes

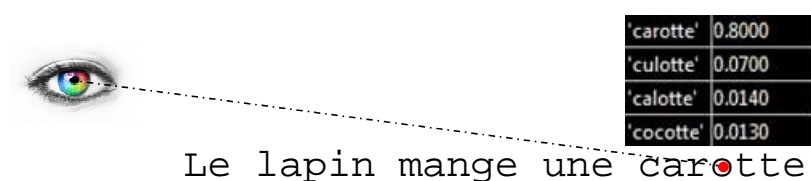
**Figure 1:** Mesure des profils d'empan visuel (PEV) pour chaque sujet.

(a) Le sujet fixe un point au centre de l'écran pendant que des groupes de 3 lettres déterminées aléatoirement sont affichés à différents emplacements du champ visuel. Les taux de reconnaissance diminuent en fonction de la distance du stimulus à la fovéa. (b) La matrice d'empan visuel (symétrique ici car seulement le champ visuel bas a été mesuré) est utile lorsqu'on cherche à simuler des patients avec perte de vision centrale. Par contre, seule la ligne horizontale centrée sur la fovéa est utile pour mesurer les limitations visuelles des sujets lors de la lecture en vision normale.



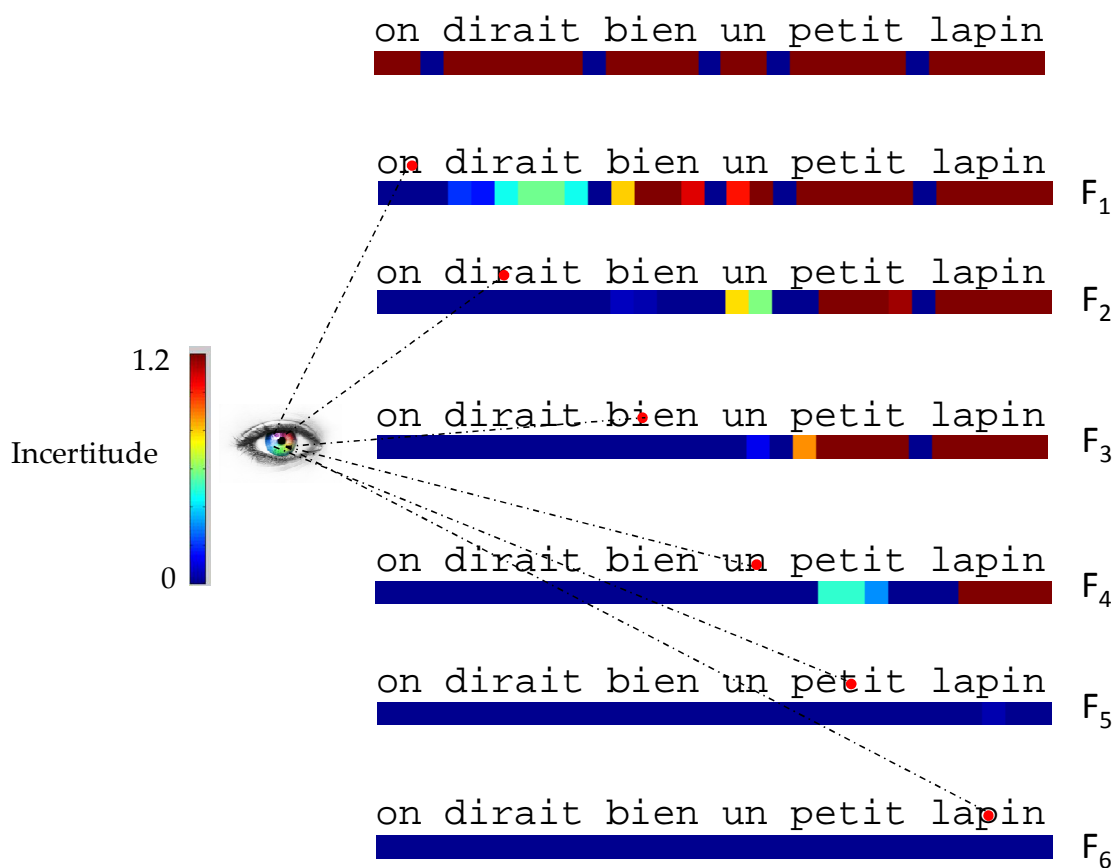
**Figure 2:** Description du fonctionnement du modèle au fur et à mesure de fixations successives.

A chaque fixation, le modèle extrait de l'information des différents mots de la phrase (ici ne sont représentés que le mot  $n$  (traité fovéalement) et le mot  $n+1$  (traité parafovéalement)). A chaque fixation, les informations de la fixation en cours sont combinées avec les informations obtenues lors des précédentes fixations. Chaque fixation permet ainsi de réduire l'incertitude sur les mots présentés dans la phrase. Les différentes valeurs correspondent aux probabilités d'activation les plus élevées du lexique. Par exemple, lorsque l'œil fixe le mot "Le", le mot est reconnu avec une probabilité  $p=0.9821$ , alors que le mot lapin traité en vision para-fovéale n'est pas complètement identifié à ce stade ( $p=0.77$ ). Notez que lors de la 4ème fixation, le mot 'une' est sauté car il a déjà été identifié ( $p=0.9760$ ) lors de la fixation précédente (sur le mot "mange").



**Figure 3:** Exemple de comportement oculomoteur prédit par le modèle.

L'incertitude pour chaque lettre L est égale à  $-\sum_{i \in [1:26]} p_i(L) * \log(p_i(L))$  où  $p_i(L)$  est la probabilité que la lettre L soit confondue avec chacune des 26 lettres de l'alphabet  $l_i$ . Avec la phrase spécifique de notre exemple, le modèle prédit une séquence de 6 fixations pour la lecture de la phrase ( $F_1, \dots, F_6$ ). Il s'agit de la première version implémentée du modèle qui ne prend en compte que la composition orthographique et la fréquence lexicale des mots. Le PEV utilisé pour cette simulation est celui montré dans la figure 1.





## **Pertinence du projet pour le BLRI**

Ce projet me semble très bien adapté aux axes « Approche neuro-computationnelle du langage » et "Pathologies du Langage" du BLRI. Il consiste en l'implémentation et la validation d'un modèle permettant de rendre compte des processus visuels, oculomoteurs et cognitifs se déroulant lors de la lecture d'un texte. L'approche est novatrice car elle utilise la théorie de l'information pour quantifier l'information extraite à chaque fixation, pour combiner l'information de deux fixations successives et pour déterminer la position et la durée de la prochaine fixation. Ce projet est également utile par sa portée appliquée aux sujets souffrant de vitesses de lecture réduites (malvoyance ou dyslexie). L'étude des comportements oculomoteurs de ces sujets permettra ainsi de définir les parts respectives des facteurs visuels, oculomoteurs et cognitifs dans la limitation des performances de lecture, ce qui pourrait permettre de définir des stratégies de rééducation adaptées aux sujets. Je pense que les résultats de ce projet pourront également être directement utiles à l'ensemble de la communauté du BLRI, via le développement d'un test de lecture hautement standardisé qui sera constitué de plusieurs centaines de phrases. Le choix de ces phrases sera rendu possible par l'utilisation de métriques de prédictibilité définies précédemment dans le projet pour la validation du modèle. Ce test pourrait ainsi être utilisé par la communauté pour mesurer avec précision les vitesses de lecture de sujets souffrant ou non de difficulté de lecture.

La collaboration entre une équipe spécialiste de la vision et de la lecture (Equipe Perception et Attention du LPC), et deux équipes spécialisées dans la linguistique et le traitement automatique du langage (Equipe Langage du LIA et Equipe Traitement Automatique du Langage Ecrit et Parlé du LIF) est naturelle et nécessaire pour ce projet multidisciplinaire mêlant psycholinguistique, psychophysique, informatique et modélisation.

Le postdoctorant potentiel (JBB) a par le passé travaillé longuement sur la mesure des performances de lecture chez les sujets sains et les patients malvoyants. JBB a également un background scientifique en informatique et mathématique qui lui a permis de développer des modèles mathématiques pour certains de ses projets de recherche, ce qui lui permettrait d'être le lien idéal entre ces trois équipes.

## **Lettre de motivation**

Intéressé depuis mes études d'ingénieur par la vision et ses aspects computationnels, j'ai obtenu une thèse en Neurosciences en 2009 (bourse ingénieur CNRS) intitulée « Psychophysique de la lecture en basse vision: Simulation de scotome, modélisation et développement d'aides visuelles », puis effectué 2 post-doctorats successifs, l'un à l'école d'Optométrie de l'université de Berkeley (2009-2013), et l'autre au Laboratoire de Psychologie Cognitive (LPC) à Marseille (2013-2015).

Mes années de recherche ont été essentiellement dédiées à la psychologie expérimentale et à l'étude psychophysique de la reconnaissance de lettres, de mots, et de la lecture chez les sujets sains et les patients sans vision centrale (pathologies de type Dégénérescence Maculaire Liée à l'Age). Mon background scientifique en informatique et modélisation m'a également permis de développer des logiciels de simulateur de déficience visuelle en temps réel, ainsi que des outils de mesure de performance de lecture.

En postdoctorat au laboratoire de Psychologie Cognitive depuis près de 2 ans, je me suis intéressé à la lecture d'un point de vue plus cognitif et moins visuel que durant ma thèse et mon précédent postdoctorat, ce qui m'a amené à considérer le développement et la validation du modèle de lecture complet présenté dans ce projet. La présentation de ce modèle à un workshop international consacré aux aspects visuels de la lecture où j'étais invité par Jonathan Grainger en Mars 2015 m'a permis de discuter avec un certain nombre de sommités du domaine (dont les développeurs des modèles de lecture les plus en pointe) de l'intérêt de cette approche probabiliste novatrice et de ses avantages comparés aux modèles existants. La version actuelle limitée du modèle vient par ailleurs d'être acceptée pour une présentation orale lors de la conférence européenne sur la vision (ECVP) qui se tiendra à Liverpool en Aout 2015. Il me semble donc important de valider cette approche le plus rapidement possible. Souhaitant continuer à appliquer mes recherches à l'aide de personnes victimes de difficultés de lecture (patients malvoyants ou sujets dyslexiques), je voudrais également utiliser ce modèle pour déterminer avec une certaine précision les causes de ces difficultés de lecture, ainsi que pour développer un test hautement standardisé de lecture.

Je compte beaucoup sur l'interaction entre un axe Neurosciences visuelles (dans lequel j'évolue depuis des années) et un axe linguistique (étude et traitement automatique du langage) pour mener à bien ce projet.

En espérant vous avoir convaincu de l'intérêt de ce projet et de ma motivation à le mener à bien.

Cordialement,

Jean-Baptiste Bernard